

#17



RECEIVED

1600

AUG 11 2003

DATE: 8/11/2003

TIME: 14:11:44

TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/734,329A

Input Set : A:\UTXC666.txt

Output Set: N:\CRF4\07302003\I734329A.raw

3 <110> APPLICANT: de CROMBRUGGHE, BENOIT.  
 4 NAKASHIMA, KAZUHISA  
 5 ZHOU, XIN  
 7 <120> TITLE OF INVENTION: MASTER BONE FORMATION TRANSCRIPTION FACTOR:  
 8 COMPOSITIONS AND METHODS OF USE  
 10 <130> FILE REFERENCE: UTXC:666  
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/734,329A  
 13 <141> CURRENT FILING DATE: 2000-11-30  
 15 <160> NUMBER OF SEQ ID NOS: 6  
 17 <170> SOFTWARE: PatentIn Ver. 2.1  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 2960  
 21 <212> TYPE: DNA  
 22 <213> ORGANISM: Mus musculus  
 24 <220> FEATURE:  
 25 <221> NAME/KEY: CDS  
 26 <222> LOCATION: (100)..(1383)  
 28 <400> SEQUENCE: 1  
 29 attctcccat tctccctccc tctcccttct ccctctccca ctggctcttc ggttctctcc 60  
 31 atctgcctga ctcttgga cccggtcccc agctcgagg atg gcg tcc tct ctg 114  
 32 Met Ala Ser Ser Leu  
 33 1 5  
 35 ctt gag gaa gaa gct cac tat ggc tcc agt ccc ctg gcc atg ctg act 162  
 36 Leu Glu Glu Glu Ala His Tyr Gly Ser Ser Pro Leu Ala Met Leu Thr  
 37 10 15 20  
 39 gca gcc tgc agc aaa ttt ggc ggc tct agc cct ctg cgg gac tca aca 210  
 40 Ala Ala Cys Ser Lys Phe Gly Gly Ser Ser Pro Leu Arg Asp Ser Thr  
 41 25 30 35  
 43 acc ctg ggg aaa gga ggc aca aag aag cca tac gct gac ctt tca gcc 258  
 44 Thr Leu Gly Lys Gly Gly Thr Lys Lys Pro Tyr Ala Asp Leu Ser Ala  
 45 40 45 50  
 47 ccc aaa acc atg ggg gac gcc tac cca gct ccc ttc tca agc acc aat 306  
 48 Pro Lys Thr Met Gly Asp Ala Tyr Pro Ala Pro Phe Ser Ser Thr Asn  
 49 55 60 65  
 51 gga ctc ctc tct cct gca ggc agt cct ccg gcc cca gcc tct ggc tat 354  
 52 Gly Leu Leu Ser Pro Ala Gly Ser Pro Pro Ala Pro Ala Ser Gly Tyr  
 53 70 75 80 85  
 55 gca aat gac tac cca ccc ttc cct cac tca ttt cct ggg ccc acc ggt 402  
 56 Ala Asn Asp Tyr Pro Pro Phe Pro His Ser Phe Pro Gly Pro Thr Gly  
 57 90 95 100  
 59 gcc caa gac cct ggg ctc cta gtg cct aag ggg cac agc tcg tct gac 450  
 60 Ala Gln Asp Pro Gly Leu Leu Val Pro Lys Gly His Ser Ser Ser Asp  
 61 105 110 115

ENTERED

## RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/734,329A

TIME: 14:11:44

Input Set : A:\UTXC666.txt

Output Set: N:\CRF4\07302003\I734329A.raw

```

63 tgc ctg cct agt gtc tac act tcc ctg gat atg act cat ccc tat ggc 498
64 Cys Leu Pro Ser Val Tyr Thr Ser Leu Asp Met Thr His Pro Tyr Gly
65      120      125      130
67 tgc tgg tac aag gca ggc atc cac gca ggc atc tca cca ggt cca ggc 546
68 Ser Trp Tyr Lys Ala Gly Ile His Ala Gly Ile Ser Pro Gly Pro Gly
69      135      140      145
71 aac aca cct act cct tgg tgg gac atg cac cct ggg ggc aac tgg cta 594
72 Asn Thr Pro Thr Pro Trp Trp Asp Met His Pro Gly Gly Asn Trp Leu
73 150      155      160      165
75 ggt ggt ggt cag ggc cag ggt gat ggg ctg caa ggg aca ctg tcc aca 642
76 Gly Gly Gly Gln Gly Gln Gly Asp Gly Leu Gln Gly Thr Leu Ser Thr
77      170      175      180
79 ggc cct gcc cag cct cca ctg aac ccc cag ctg cct act tac cca tct 690
80 Gly Pro Ala Gln Pro Pro Leu Asn Pro Gln Leu Pro Thr Tyr Pro Ser
81      185      190      195
83 gac ttt gct ccc ctt aac cca gct ccc tac cca gcg ccc cac ctc ttg 738
84 Asp Phe Ala Pro Leu Asn Pro Ala Pro Tyr Pro Ala Pro His Leu Leu
85      200      205      210
87 caa cca ggg ccc cag cat gtc cta ccc caa gat gtc tat aag ccc aag 786
88 Gln Pro Gly Pro Gln His Val Leu Pro Gln Asp Val Tyr Lys Pro Lys
89      215      220      225
91 gcg gtt ggc aat agt ggg caa ctg gag ggg agt ggt gca gcc aaa ccc 834
92 Ala Val Gly Asn Ser Gly Gln Leu Glu Gly Ser Gly Ala Ala Lys Pro
93 230      235      240      245
95 cct cgg ggt gct ggc aca ggg ggc agc ggt gga tat gcg ggc agt ggg 882
96 Pro Arg Gly Ala Gly Thr Gly Gly Ser Gly Gly Tyr Ala Gly Ser Gly
97      250      255      260
99 gca ggg cgt tct acc tgc gac tgc ccc aac tgt cag gag cta gag cgg 930
100 Ala Gly Arg Ser Thr Cys Asp Cys Pro Asn Cys Gln Glu Leu Glu Arg
101      265      270      275
103 ctc ggg gca gca gcg gct ggg ctg agg aag aag ccc att cac agc tgc 978
104 Leu Gly Ala Ala Ala Ala Gly Leu Arg Lys Lys Pro Ile His Ser Cys
105      280      285      290
107 cac atc cct ggg tgc ggc aag gtg tac ggc aag gct tcg cat ctg aaa 1026
108 His Ile Pro Gly Cys Gly Lys Val Tyr Gly Lys Ala Ser His Leu Lys
109      295      300      305
111 gcc cac ttg cgc tgg cac act ggc gag agg cct ttc gtc tgc aac tgg 1074
112 Ala His Leu Arg Trp His Thr Gly Glu Arg Pro Phe Val Cys Asn Trp
113 310      315      320      325
115 ctt ttc tgc ggc aag agg ttc act cgc tct gac gag ctg gag cgc cac 1122
116 Leu Phe Cys Gly Lys Arg Phe Thr Arg Ser Asp Glu Leu Glu Arg His
117      330      335      340
119 gtg cgc act cac acc cgg gag aag aag ttc act tgc ctg ctc tgt tcc 1170
120 Val Arg Thr His Thr Arg Glu Lys Lys Phe Thr Cys Leu Leu Cys Ser
121      345      350      355
123 aag cgc ttt acc aga agc gac cac ttg agc aaa cat cag cgc acc cac 1218
124 Lys Arg Phe Thr Arg Ser Asp His Leu Ser Lys His Gln Arg Thr His
125      360      365      370
127 ggg gag cca ggc ccg gga ccg ccc cca agt ggc cct aag gag ctg ggg 1266

```

## RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/734,329A

TIME: 14:11:44

Input Set : A:\UTXC666.txt

Output Set: N:\CRF4\07302003\I734329A.raw

```

128 Gly Glu Pro Gly Pro Gly Pro Pro Pro Ser Gly Pro Lys Glu Leu Gly
129      375                      380                      385
131 gag ggt cgc agc gtc ggg gaa gaa gaa gcc aat cag ccg ccc cga tct 1314
132 Glu Gly Arg Ser Val Gly Glu Glu Glu Ala Asn Gln Pro Pro Arg Ser
133 390                      395                      400                      405
135 tcc act tcg cct gca ccc cca gaa aaa gcc cac gga ggc agc cca gag 1362
136 Ser Thr Ser Pro Ala Pro Pro Glu Lys Ala His Gly Gly Ser Pro Glu
137      410                      415                      420
139 cag agc aac ctg cta gag atc tgagccgggt agaggaaggt ctccagctcc 1413
140 Gln Ser Asn Leu Leu Glu Ile
141      425
143 agggctcctct tgccaggctc tcttggcgtg ctggacccat tggttgcccc tcgctctctc 1473
145 ctattgcatg ctatactctg ggggctctct ctgttcccct aggctatctc cttgcatgtc 1533
147 tcctcagttc ttctctcttt gtcaagagtc ttagccaaac tcctctcagg cctttgccag 1593
149 tgcctagttc ctatgctccg acctcctcaa ctttttcttc tctgcccctg ttcttcacag 1653
151 cttccatctg gcctcacatc attttctcat taactcgttg ccatctaatac tttctgcttc 1713
153 ccaatcctat ttgccgtttt cccgaagctt ccaggctgtc gcctcgattc cccccacct 1773
155 ttcgtcttcc tgagctttgt gttttctttt tttaaacaaa cacgatgatg atgatgatga 1833
157 tgatgataat ttattgcccc ctgggtgttct tcattaggaa ccagagttaa ggagattggg 1893
159 gttagtaacc tggccgggag cagagtgcc aagaaggggga agtccaatgg ggatctgatc 1953
161 ccaaagatgg ggtgaccca gggtcaggga ggctgcccc agccttgagt acttaacccc 2013
163 tatgcgccag gagtaaagaa tagtaatagt aataataata ataattctat ttatctaagt 2073
165 tatgatgacg ggtcaggtag agtgagctgg agagggaag ggattctccc cgcccccaag 2133
167 gaaattctag tcaaatgcat ctctgtatag acaaatgata gtggagacct tgctcgtaga 2193
169 tttctatcct cgaggtctcc gagagtttct ttttcagttg agttttgggt tgttcggcct 2253
171 ctttttagagt ttctgtgggt gtctctctgt taggcagtca ctaagatccc cagcccagc 2313
173 cagaaagctg tgaacttca agtcctatgg cggggaggac tggaatgtac cccagtcctc 2373
175 tcgacccgac tgcagatcag gttcctcccc tgatcctctt ctcataccct gtgacctcac 2433
177 caggttatcc ccttgctgctc atggttacag agagcttgca gctgccatct taaacgtgct 2493
179 ctttggggga gagccacct aacaggagga ttttggtttg gaggtgcccc tcctgaaaaa 2553
181 gtaggtgggc aaaggctttc tctgggatca aattcaaata aatcaagtat ttattgaatg 2613
183 cttaatatgt gcaaggcctg gtgcctagaa gccacgagaa agaatttata acaggacaga 2673
185 agtccctaaa ctaaacatcc acaggccccc aatctaggag gtttcaactcc attccagtga 2733
187 cttttaaagc cgctttgtgc ctttgaaatg cctttcctga gatttttgga tcttcctggt 2793
189 ctgtcccctg ctccttctag gcctcaagat aaagggtaaa gccatggagt ctgggaagag 2853
191 cataacgtcg ttgacgggat cgtccctttg tggaatcttt cttttttttt taatttaata 2913
193 aataaaagtt cgatttcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 2960
196 <210> SEQ ID NO: 2
197 <211> LENGTH: 428
198 <212> TYPE: PRT
199 <213> ORGANISM: Mus musculus
201 <400> SEQUENCE: 2
202 Met Ala Ser Ser Leu Leu Glu Glu Glu Ala His Tyr Gly Ser Ser Pro
203 1 5 10 15
205 Leu Ala Met Leu Thr Ala Ala Cys Ser Lys Phe Gly Gly Ser Ser Pro
206 20 25 30
208 Leu Arg Asp Ser Thr Thr Leu Gly Lys Gly Gly Thr Lys Lys Pro Tyr
209 35 40 45
211 Ala Asp Leu Ser Ala Pro Lys Thr Met Gly Asp Ala Tyr Pro Ala Pro

```

## RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/734,329A

TIME: 14:11:44

Input Set : A:\UTXC666.txt

Output Set: N:\CRF4\07302003\I734329A.raw

```

212      50      55      60
214 Phe Ser Ser Thr Asn Gly Leu Leu Ser Pro Ala Gly Ser Pro Pro Ala
215 65      70      75      80
217 Pro Ala Ser Gly Tyr Ala Asn Asp Tyr Pro Pro Phe Pro His Ser Phe
218      85      90      95
220 Pro Gly Pro Thr Gly Ala Gln Asp Pro Gly Leu Leu Val Pro Lys Gly
221      100      105      110
223 His Ser Ser Ser Asp Cys Leu Pro Ser Val Tyr Thr Ser Leu Asp Met
224      115      120      125
226 Thr His Pro Tyr Gly Ser Trp Tyr Lys Ala Gly Ile His Ala Gly Ile
227      130      135      140
229 Ser Pro Gly Pro Gly Asn Thr Pro Thr Pro Trp Trp Asp Met His Pro
230 145      150      155      160
232 Gly Gly Asn Trp Leu Gly Gly Gly Gln Gly Gln Gly Asp Gly Leu Gln
233      165      170      175
235 Gly Thr Leu Ser Thr Gly Pro Ala Gln Pro Pro Leu Asn Pro Gln Leu
236      180      185      190
238 Pro Thr Tyr Pro Ser Asp Phe Ala Pro Leu Asn Pro Ala Pro Tyr Pro
239      195      200      205
241 Ala Pro His Leu Leu Gln Pro Gly Pro Gln His Val Leu Pro Gln Asp
242      210      215      220
244 Val Tyr Lys Pro Lys Ala Val Gly Asn Ser Gly Gln Leu Glu Gly Ser
245 225      230      235      240
247 Gly Ala Ala Lys Pro Pro Arg Gly Ala Gly Thr Gly Gly Ser Gly Gly
248      245      250      255
250 Tyr Ala Gly Ser Gly Ala Gly Arg Ser Thr Cys Asp Cys Pro Asn Cys
251      260      265      270
253 Gln Glu Leu Glu Arg Leu Gly Ala Ala Ala Ala Gly Leu Arg Lys Lys
254      275      280      285
256 Pro Ile His Ser Cys His Ile Pro Gly Cys Gly Lys Val Tyr Gly Lys
257      290      295      300
259 Ala Ser His Leu Lys Ala His Leu Arg Trp His Thr Gly Glu Arg Pro
260 305      310      315      320
262 Phe Val Cys Asn Trp Leu Phe Cys Gly Lys Arg Phe Thr Arg Ser Asp
263      325      330      335
265 Glu Leu Glu Arg His Val Arg Thr His Thr Arg Glu Lys Lys Phe Thr
266      340      345      350
268 Cys Leu Leu Cys Ser Lys Arg Phe Thr Arg Ser Asp His Leu Ser Lys
269      355      360      365
271 His Gln Arg Thr His Gly Glu Pro Gly Pro Gly Pro Pro Pro Ser Gly
272      370      375      380
274 Pro Lys Glu Leu Gly Glu Gly Arg Ser Val Gly Glu Glu Glu Ala Asn
275 385      390      395      400
277 Gln Pro Pro Arg Ser Ser Thr Ser Pro Ala Pro Pro Glu Lys Ala His
278      405      410      415
280 Gly Gly Ser Pro Glu Gln Ser Asn Leu Leu Glu Ile
281      420      425
285 <210> SEQ ID NO: 3
286 <211> LENGTH: 14

```

## RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/734,329A

TIME: 14:11:44

Input Set : A:\UTXC666.txt

Output Set: N:\CRF4\07302003\I734329A.raw

```

287 <212> TYPE: PRT
288 <213> ORGANISM: Mus musculus
290 <400> SEQUENCE: 3
291 Ala His Gly Gly Ser Pro Glu Gln Ser Asn Leu Leu Glu Ile
292   1           5           10
295 <210> SEQ ID NO: 4
296 <211> LENGTH: 85
297 <212> TYPE: PRT
298 <213> ORGANISM: Mus musculus
300 <400> SEQUENCE: 4
301 Ile His Ser Cys His Ile Pro Gly Cys Gly Lys Val Tyr Gly Lys Ala
302   1           5           10           15
304 Ser His Leu Lys Ala His Leu Arg Trp His Thr Gly Glu Arg Pro Phe
305           20           25           30
307 Val Cys Asn Trp Leu Phe Cys Gly Lys Arg Phe Thr Arg Ser Asp Glu
308           35           40           45
310 Leu Glu Arg His Val Arg Thr His Thr Arg Glu Lys Lys Phe Thr Cys
311           50           55           60
313 Leu Leu Cys Ser Lys Arg Phe Thr Arg Ser Asp His Leu Ser Lys His
314   65           70           75           80
316 Gln Arg Thr His Gly
317           85
320 <210> SEQ ID NO: 5
321 <211> LENGTH: 244
322 <212> TYPE: PRT
323 <213> ORGANISM: Mus musculus
325 <400> SEQUENCE: 5
326 Phe Gly Gly Ser Ser Pro Leu Arg Asp Ser Thr Thr Leu Gly Lys Gly
327   1           5           10           15
329 Gly Thr Lys Lys Pro Tyr Ala Asp Leu Ser Ala Pro Lys Thr Met Gly
330           20           25           30
332 Asp Ala Tyr Pro Ala Pro Phe Ser Ser Thr Asn Gly Leu Leu Ser Pro
333           35           40           45
335 Ala Gly Ser Pro Pro Ala Pro Ala Ser Gly Tyr Ala Asn Asp Tyr Pro
336           50           55           60
338 Pro Phe Pro His Ser Phe Pro Gly Pro Thr Gly Ala Gln Asp Pro Gly
339   65           70           75           80
341 Leu Leu Val Pro Lys Gly His Ser Ser Ser Asp Cys Leu Pro Ser Val
342           85           90           95
344 Tyr Thr Ser Leu Asp Met Thr His Pro Tyr Gly Ser Trp Tyr Lys Ala
345           100          105          110
347 Gly Ile His Ala Gly Ile Ser Pro Gly Pro Gly Asn Thr Pro Thr Pro
348           115          120          125
350 Trp Trp Asp Met His Pro Gly Gly Asn Trp Leu Gly Gly Gly Gln Gly
351           130          135          140
353 Gln Gly Asp Gly Leu Gln Gly Thr Leu Ser Thr Gly Pro Ala Gln Pro
354   145          150          155          160
356 Pro Leu Asn Pro Gln Leu Pro Thr Tyr Pro Ser Asp Phe Ala Pro Leu
357           165          170          175

```

**VERIFICATION SUMMARY**

DATE: 07/30/2003

PATENT APPLICATION: US/09/734,329A

TIME: 14:11:45

Input Set : A:\UTXC666.txt

Output Set: N:\CRF4\07302003\I734329A.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number